W2. Gaussian Distribution

Phuong Anh Trinh - 11200417

August 2022

1 Exercise 1a: Proof that Gaussian distribution is normalized

$p(x|\mu,\sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}}.e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

To prove for normalized, we have to show that:

$\int\limits_{-\infty}^{\infty} p(x|\mu,\sigma^2)dx = 1$

$\Leftrightarrow \dfrac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} = 1$

Assume $\mu = 0$

$\int\limits_{-\infty}^{\infty} e^{\frac{-x^2}{2\sigma^2}} dx = \sqrt{2\pi\sigma^2} \qquad (1)$

Let $I = \int\limits_{-\infty}^{\infty} e^{\frac{-x^2}{2\sigma^2}} dx$

$I^2 = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{\frac{-(x^2+y^2)}{2\sigma^2}} dxdy \qquad (2)$

Let $x = r.sin\theta, y = r.cos\theta$

$\dfrac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \dfrac{\partial(x)}{\partial(r)} & \dfrac{\partial(x)}{\partial(\theta)} \\ \dfrac{\partial(y)}{\partial(r)} & \dfrac{\partial(y)}{\partial(\theta)} \end{vmatrix} = \begin{vmatrix} cos\theta & -r.sin\theta \\ sin\theta & r.cos\theta \end{vmatrix}$

$= r.cos^2\theta + r.sin^2\theta$

$= r.(sin^\theta + cos^2\theta = r$

Apply to (2), we have:

$I^2 = \int\limits_0^{2\pi}\int\limits_0^{\infty} e^{\frac{-r^2}{2\sigma^2}} rdrd\theta$

$= 2\pi \int\limits_0^{\infty} e^{\frac{-r^2}{2\sigma^2}} rdr$

$= 2\pi \int\limits_0^{\infty} e^{\frac{-u}{2\sigma^2}} \dfrac{1}{2}du \qquad (\text{Let } r^2 = u \Rightarrow rdr = \dfrac{1}{2}du)$

$= 2\pi\sigma^2$

$\Rightarrow I = \sqrt{2\pi\sigma^2}$ satisfy (1) $\Rightarrow$ proved

2 Exercise 1b: Proof that Expectation of Gaussian distribution is μ

$E(X) = \int\limits_{-\infty}^{\infty} x.f(x)dx$

$$\Leftrightarrow E(X) = \int\limits_{-\infty}^{\infty} x.\frac{1}{\sqrt{2\pi\sigma^2}}.e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{1}$$

Let $\frac{x-\mu}{\sigma\sqrt{2}} = t$

$\Rightarrow dx = \sigma\sqrt{2}dt$

Apply to (1), we have:

$E(X) = \int\limits_{-\infty}^{\infty} \frac{\sigma t\sqrt{2}+\mu}{\sigma\sqrt{2\pi}}.\sigma\sqrt{2}.e^{-t^2}dt$

$= \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (\sigma t\sqrt{2}+\mu).e^{-t^2}dt$

$= \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (\sigma t\sqrt{2}.e^{-t^2} + \mu.e^{-t^2})dt$

$= \frac{1}{\pi}(\sigma\sqrt{2} \int\limits_{-\infty}^{\infty} t.e^{-t^2}dt + \mu \int\limits_{-\infty}^{\infty} e^{-t^2}dt)$

$= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}} = \mu \Rightarrow$ proved

3 Exercise 1c: Proof that Variance of Gaussian distribution is σ^2

$Var(X) = \int\limits_{-\infty}^{\infty} x^2.f(x)dx - (E(X))^2$

$$\Leftrightarrow \int\limits_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}} - \mu^2 \tag{1}$$

Let $\frac{x-\mu}{\sigma\sqrt{2}} = t$

$\Rightarrow dx = \sigma\sqrt{2}dt$

Apply to (1), we have:

$\int\limits_{-\infty}^{\infty} (\sigma t\sqrt{2}+\mu)^2.\frac{\sigma\sqrt{2}}{\sqrt{2\pi\sigma^2}}.e^{-t^2}dt - \mu^2$

$= \frac{1}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} (2\sigma^2t^2 + 2\sqrt{2}\sigma t\mu + \mu^2).e^{-t^2}dt - \mu^2$

$= \frac{1}{\sqrt{\pi}}(\int\limits_{-\infty}^{\infty} 2\sigma^2t^2.e^{-t^2}dt + \int\limits_{-\infty}^{\infty} 2\sqrt{2}\sigma t\mu.e^{-t^2}dt + \int\limits_{-\infty}^{\infty} \mu^2.e^{-t^2}dt) - \mu^2$

$= \frac{2\sigma^2}{\sqrt{\pi}}(\int\limits_{-\infty}^{\infty} t^2dt + \int\limits_{-\infty}^{\infty} e^{-t^2}dt) + 0 + \frac{\mu^2\sqrt{\pi}}{\sqrt{\pi}} - \mu^2$

$= \frac{2\sigma^2.\sqrt{\pi}}{2\sqrt{\pi}} = \sigma^2 \Rightarrow$ proved

4 Exercise 1d: Mutivariate Gaussian is normalized

For a D-dimension vector x, the multivariate Gaussian distribution has the form

$$p(x|\mu,\sigma^2) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}}.e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

Set

$$\triangle^2 = (x-\mu)^T\sum^{-1}(x-\mu)$$
$$= -\frac{1}{2}x^T\sum^{-1}x + x^T\sum^{-1}\mu + const$$

Because Σ is a real, symmetric matrix, its eigenvalues will be real and its eigenvectors form an orthonormal set

$$\Sigma = \sum_{i=1}^{D}\lambda_i u_i u_i^T \Rightarrow \Sigma^{-1} = \sum_{i=1}^{D}\frac{1}{\lambda_i}u_i u_i^T$$

So that

$$\Delta^2 = (x-\mu)^T\Sigma^{-1}(x-\mu) = \sum_{i=1}^{D}\frac{1}{\lambda_i}(x-\mu)^T u_i u_i^T(x-\mu)$$
$$= \sum_{i=1}^{D}\frac{y_i^2}{\lambda_i}, \text{with} y_i = u_i^T(x-\mu)$$

$$|\Sigma|^{1/2} = \prod_{j=1}^{D}\lambda_j^{1/2}$$

$$p(y) = \prod_{j=1}^{D}\frac{1}{2\pi\lambda_j}^{1/2}.e^{-\frac{y_j^2}{2\lambda_j}}$$

$$\Rightarrow \int_{-\infty}^{\infty}p(y)dy = \prod_{j=1}^{D}\int_{-\infty}^{\infty}\frac{1}{2\pi\lambda_j}^{1/2}.e^{-\frac{y_j^2}{2\lambda_j}}dy_j = 1 \rightarrow \text{proved}$$

5 Exercise 2a: Condition Gaussian distribution

Suppose x is a D-dimension vector with Gaussian distribution and we partition x into two disjoint subsets x_a and x_b

$$x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$$

Mean vector μ

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

Covariance matrix $\sum$

$$\sum = \begin{pmatrix} \sum_{aa} & \sum_{ab} \\ \sum_{ba} & \sum_{bb} \end{pmatrix} \Rightarrow A = \sum^{-} 1 = \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix}$$

$\sum$ is symmetric so $\sum_{aa}$ and $\sum_{bb}$ are symmetric while $\sum_{ab} = \sum_{ba}^T$

We have

$-\frac{1}{2}(x-\mu)^T \sum^{-1}(x-\mu)$

$= -\frac{1}{2}(x-\mu)^T A(x-\mu)$

$= -\frac{1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \frac{1}{2}(x_a-\mu_a)^T A_{ab}(x_b-\mu_b)$

$\quad -\frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a) - \frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b)$

$= -\frac{1}{2}x_a^T A_{aa}^{-1} x_a + x_a^T(A_{aa}\mu_a - A_{ab}(x_b-\mu_b)) + const$

Compare with Gaussian distribution

$\triangle^2 = -\frac{1}{2}x^T \sum^{-1} x + x^T \sum^{-1} \mu + const$

$\sum_{a|b} = A_{aa}^{-1}$

$\mu_{a|b} = \sum_{a|b}(Aaa\mu_a - A_{ab}(x_b-\mu_b)) = \mu_a - A_{aa}^{-1}A_{ab}(x_b-\mu_b)$

By using Schur complement

$A_{aa} = (\sum_{aa} - \sum_{ab}\sum_{bb}^{-1}\sum_{ba})^{-1}$

$A_{ab} = -(\sum_{aa} - \sum_{ab}\sum_{bb}^{-1}\sum_{ba})^{-1}\sum_{ab}\sum bb^{-1}$

As a result

$\mu_{a|b} = \mu_a + \sum_{ab}\sum_{bb}^{-1}(x_b-\mu_b)$

$\sum_{a|b} = \sum_{aa} - \sum ab \sum bb^{-1} \sum_{ba}$

$\Rightarrow p(x_a|x_b) = N(x_{a|b}|\mu_{a|b}, \sum_{a|b})$

6 Exercise 2b: Marginal Gaussian distribution

The marginal distribution given by

$p(x_a) = \int p(x_a, x_b)dx_b$

We have

$-\frac{1}{2}x_b^T A_{bb} x_b + x_b^T m = -\frac{1}{2}(x_b - A_{aa}^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1} + \frac{1}{2}m^T A_{bb}^{-1} m$

Define $m = A_{bb}\mu_b - A_{ba}(x_a-\mu_a)$

Integrate over unnormalized Gaussian

$\int e^{-\frac{1}{2}(x_b - A_{bb}^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1}m)} dx_b$

The remaining term

$-\frac{1}{2}x_a^T(A_{aa} - A_{ab}A_{bb}^{-1}A_{ba})x_a + x_a^T(A_{aa} - A_{ab}A_{bb}^{-1}A_{ba})^{-}1\mu_a + const$

Similarly, we have

$$E(x_a) = \mu_a \qquad cov(x_a) = \sum_{aa}$$

$$\Rightarrow p(x_a) = N(x_a|\mu_a, \sum_{aa})$$